

# SEQUENCE LISTING

<110> Nilles, Matthew L.  
 Matson, Jyl S.  
 <120> YERSINIA SPECIES COMPOSITIONS  
 <130> 3128-6046US  
 <150> US 60/444,076  
 <151> 2003-01-31  
 <160> 20  
 <170> PatentIn version 3.2  
 <210> 1  
 <211> 87  
 <212> PRT  
 <213> Yersinia pestis  
 <220>  
 <223> amino acid sequence of YscF  
 <220>  
 <223> sequence can be found at MedLine accession number NP\_857921.1  
 <400> 1  
 Met Ser Asn Phe Ser Gly Phe Thr Lys Gly Thr Asp Ile Ala Asp Leu  
 1 5 10 15  
 Asp Ala Val Ala Gln Thr Leu Lys Lys Pro Ala Asp Asp Ala Asn Lys  
 20 25 30  
 Ala Val Asn Asp Ser Ile Ala Ala Leu Lys Asp Lys Pro Asp Asn Pro  
 35 40 45  
 Ala Leu Leu Ala Asp Leu Gln His Ser Ile Asn Lys Trp Ser Val Ile  
 50 55 60  
 Tyr Asn Ile Asn Ser Thr Ile Val Arg Ser Met Lys Asp Leu Met Gln  
 65 70 75 80  
 Gly Ile Leu Gln Lys Phe Pro  
 85  
 <210> 2  
 <211> 264  
 <212> DNA  
 <213> Yersinia pestis  
 <220>  
 <223> nucleic acids encoding YscF  
 <220>  
 <223> sequence can be found at MedLine accession number NL\_004839  
 <400> 2

atgagtaact tctctggatt tacgaaagga accgatatcg cagacttaga tgcggtggct 60  
 caaacgctca agaagccagc agacgatgca aacaaagcgg ttaatgactc gatagcagca 120  
 ttgaaagata agcctgacaa cccggcgcta ctgtctgact tacaacattc aattaataaa 180  
 tggtcggtaa ttacaatat aaactcaacc atagtctgta gcatgaaaga cttaatgcaa 240  
 ggcatcctac agaagttccc ataa 264

<210> 3  
 <211> 264  
 <212> DNA  
 <213> *Yersinia enterocolitica*

<220>  
 <223> nucleic acid encoding YscF from *Y. enterocolitica*

<220>  
 <223> sequence can be found at MedLine accession number NC\_005017

<400> 3  
 atgagtaatt tctctgggtt tgcaaaagga accgatatca cagacttaga tgcggtggct 60  
 caaacgctca agaagccagc agacggcgca aacaaagcgg ttaatgactc gatagcagcg 120  
 ttgaaagaga cgcctgacaa cccggcgcta ctgtctgact tacaacattc aattaataaa 180  
 tggtcggtaa ttacaatat aagctcaacc atagtctgta gcatgaaaga cttaatgcaa 240  
 ggcatcctac agaagttccc ataa 264

<210> 4  
 <211> 87  
 <212> PRT  
 <213> *Yersinia enterocolitica*

<220>  
 <223> amino acid sequence of YscF from *Y. enterocolitica*

<220>  
 <223> sequence can be found at MedLine accession number NP\_863538.1

<400> 4

Met Ser Asn Phe Ser Gly Phe Ala Lys Gly Thr Asp Ile Thr Asp Leu  
 1 5 10 15  
 Asp Ala Val Ala Gln Thr Leu Lys Lys Pro Ala Asp Asp Ala Asn Lys  
 20 25 30  
 Ala Val Asn Asp Ser Ile Ala Ala Leu Lys Asp Thr Pro Asp Asn Pro  
 35 40 45  
 Ala Leu Leu Ala Asp Leu Gln His Ser Ile Asn Lys Trp Ser Val Ile  
 50 55 60

Tyr Asn Met Ser Ser Thr Ile Val Arg Ser Met Lys Asp Leu Met Gln  
65 70 75 80

Gly Ile Leu Gln Lys Phe Pro  
85

<210> 5  
<211> 80  
<212> PRT  
<213> Salmonella enterica

<220>  
<223> amino acid sequence of PrgI

<220>  
<223> sequence can be found at MedLine accession number CAD05980.1

<400> 5

Met Pro Thr Ser Trp Ser Gly Tyr Leu Asp Glu Val Ser Ala Lys Phe  
1 5 10 15

Asp Lys Gly Val Asp Asn Leu Gln Thr Gln Val Thr Glu Ala Leu Asp  
20 25 30

Lys Leu Ala Ala Lys Pro Ser Asp Pro Ala Leu Leu Ala Ala Tyr Gln  
35 40 45

Ser Lys Leu Ser Glu Tyr Asn Leu Tyr Arg Asn Ala Gln Ser Asn Thr  
50 55 60

Val Lys Val Phe Lys Asp Ile Asp Ala Ala Ile Ile Gln Asn Phe Arg  
65 70 75 80

<210> 6  
<211> 243  
<212> DNA  
<213> Salmonella enterica

<220>  
<223> sequence can be found at MedLine accession number AL627276

<220>  
<223> nucleic acid encoding PrgI

<400> 6  
atgccaaacat cttggtcagg ctatctggat gaagtttcag caaaatttga taaggcgctt 60  
gataatctac aaacgcaggt aacagaggcg ctggataaat tagcagcaaa accctccgat 120  
ccggcgctac tggcggcgta tcagagtaag ctctcggaat ataacttgta ccgtaacgcg 180  
caatcgaaca cggtaaaaagt cttaagat attgatgctg ccattattca gaacttcctg 240  
taa 243

<210> 7  
 <211> 83  
 <212> PRT  
 <213> Shigella flexneri  
  
 <220>  
 <223> amino acid sequence of MxiH  
  
 <220>  
 <223> sequence can be found at MedLine accession number NP\_858270.1  
  
 <400> 7

Met Ser Val Thr Val Pro Asn Asp Asp Trp Thr Leu Ser Ser Leu Ser  
 1 5 10 15  
  
 Glu Thr Phe Asp Asp Gly Thr Gln Thr Leu Gln Gly Glu Leu Thr Leu  
 20 25 30  
  
 Ala Leu Asp Lys Leu Ala Lys Asn Pro Ser Asn Pro Gln Leu Leu Ala  
 35 40 45  
  
 Glu Tyr Gln Ser Lys Leu Ser Glu Tyr Thr Leu Tyr Arg Asn Ala Gln  
 50 55 60  
  
 Ser Asn Thr Val Lys Val Ile Lys Asp Val Asp Ala Ala Ile Ile Gln  
 65 70 75 80  
  
 Asn Phe Arg

<210> 8  
 <211> 252  
 <212> DNA  
 <213> Shigella flexneri  
  
 <220>  
 <223> nucleic acid encoding MxiH  
  
 <220>  
 <223> sequence can be found at MedLine accession number Nc\_004851  
  
 <400> 8  
 atgagtgtta cagtaccgaa tgatgattgg acattgagtt cattatctga aacttttgat 60  
 gatggaactc aaacattaca aggtgaacta acattggcac tagataaatt agctaaaaat 120  
 ccttcgaatc cacagttgct ggctgaatac caaagtaaat tatctgaata tacattatat 180  
 aggaacgcgc aatccaatac agtgaaagtg attaaggatg ttgatgctgc aattattcaa 240  
 aacttcagat aa 252  
  
 <210> 9  
 <211> 73  
 <212> PRT  
 <213> Escherichia coli

<220>  
 <223> amino acid sequence of EscF  
 <220>  
 <223> sequence can be found at MedLine accession number NP\_312579.1  
 <400> 9

Met Asn Leu Ser Glu Ile Thr Gln Gln Met Gly Glu Val Gly Lys Thr  
 1 5 10 15

Leu Ser Asp Ser Val Pro Glu Leu Leu Asn Ser Thr Asp Leu Val Asn  
 20 25 30

Asp Pro Glu Lys Met Leu Glu Leu Gln Phe Ala Val Gln Gln Tyr Ser  
 35 40 45

Ala Tyr Val Asn Val Glu Ser Gly Met Leu Lys Thr Ile Lys Asp Leu  
 50 55 60

Val Ser Thr Ile Ser Asn Arg Ser Phe  
 65 70

<210> 10  
 <211> 222  
 <212> DNA  
 <213> Escherichia coli

<220>  
 <223> nucleic acid encoding EscF

<220>  
 <223> sequence can be found at MedLine accession number NC\_002695

<400> 10  
 atgaatttat ctgaaattac tcaacaaatg ggtgaagtag gtaaaacgct gacgcattct 60  
 gtgccagagt tacttaatag caccgatttg gttaatgacc ctgaaaaaat gttagagttg 120  
 cagtttgcgg tttagcaata ttctgcttat gttaacgtag aaagtggaat gttgaaaacg 180  
 ataaaaatgc tgggtctcaac catttctaac cgtagttttt aa 222

<210> 11  
 <211> 240  
 <212> DNA  
 <213> Yersinia pestis

<220>  
 <223> nucleic acid encoding HT-YscF; an example of a recombinant YscF

<220>  
 <221> misc\_feature  
 <222> (214)..(237)  
 <223> His-tag sequence

```

<400> 11
atgagtaact tctctggatt tacgaaagga accgatgatg caaacaaga ggttaatgac 60
tcgatagcag cattgaaaga taagcctgac aaccggcgc tacttgctga cttacaacat 120
tcaattaata aatggtcggg aatttacaat ataaactcaa ccatagtctg tagcatgaaa 180
gacttaatgc aaggcatcct acagaagttc ccactcgagc accaccacca ccaccactga 240

```

```

<210> 12
<211> 95
<212> PRT
<213> Yersinia pestis

```

```

<220>
<223> amino acid sequence of HT-YscF; an example of a recombinant YscF

```

```

<220>
<221> misc_feature
<222> (88)..(95)
<223> His-tag

```

```

<400> 12

```

```

Met Ser Asn Phe Ser Gly Phe Thr Lys Gly Thr Asp Ile Ala Asp Leu
1      5      10      15

```

```

Asp Ala Val Ala Gln Thr Leu Lys Lys Pro Ala Asp Asp Ala Asn Lys
20      25      30

```

```

Ala Val Asn Asp Ser Ile Ala Ala Leu Lys Asp Lys Pro Asp Asn Pro
35      40      45

```

```

Ala Leu Leu Ala Asp Leu Gln His Ser Ile Asn Lys Trp Ser Val Ile
50      55      60

```

```

Tyr Asn Ile Asn Ser Thr Ile Val Arg Ser Met Lys Asp Leu Met Gln
65      70      75      80

```

```

Gly Ile Leu Gln Lys Phe Pro Leu Glu His His His His His His
85      90      95

```

```

<210> 13
<211> 5537
<212> DNA
<213> Artificial Sequence

```

```

<220>
<223> pMJ119

```

```

<220>
<221> misc_feature
<222> (563)..(1375)
<223> kan encoded on complementary strand

```

```

<220>
<221> misc_feature

```

<222> (3518)..(4597)  
 <223> LacI encoded on complementary strand

<220>  
 <221> promoter  
 <222> (4983)..(5000)  
 <223> T7 promoter

<220>  
 <221> misc\_feature  
 <222> (5001)..(5001)  
 <223> T7 transcription start

<220>  
 <221> misc\_feature  
 <222> (5072)..(5104)  
 <223> T7 Tag

<220>  
 <221> CDS  
 <222> (5114)..(5401)  
 <223> HT-YscF

<220>  
 <221> misc\_feature  
 <222> (5381)..(5398)  
 <223> His-tag sequence

<220>  
 <221> terminator  
 <222> (5466)..(5512)  
 <223> T7 terminator

```

<400> 13
tggcgaatgg gacgcgccct gtagcggcgc attaagcgcg gcgggtgtgg tggttacgcg      60
cagcgtgacc gctacacttg ccagcgccct agcgcccgct ctttcgctt tcttcacctc      120
ctttctcgcc acgttcgcgg gctttcccg tcaagctcta aatcgggggc tccttttagg      180
gttcggaatt agtgctttac ggcacctcga ccccaaaaaa cttgattagg gtgatgggtc      240
acgtagtggg ccatacgccct gatagacggt ttttcgccct ttgacgttgg agtccacggt      300
ctttaatagt ggactcttgg tccaaactgg aacaacactc aacctatct cggtctattc      360
ttttgattta taagggaatt tgccgatttc ggccatttgg ttaaaaaatg agctgattta      420
acaaaaattt aacgcgaatt ttaacaaat attaacgttt acaatttcag gtggcacttt      480
tcggggaat gtgcgcggaa cccctatttg tttatttttc taaatacatt caaatatgta      540
tcggtcatg aattaatttc tagaaaaact catcgagcat caaatgaac tgcaatttat      600
tcatatcagg attatcaata ccataTTTTT gaaaagccg tttctgtaat gaaggagaaa      660
actcaccgag gcagttccat aggatggcaa gatcctgcta tcggtctgcg attccgactc      720
gtccaacatc aatacaacct attaatTTTC cctcgtcaaa aataagggtta tcaagtgaga      780
aatcaccatg agtgacgact gaatccggtg agaattggcaa aagtttatgc atttctttcc      840
agacttggtc aacaggccag ccattacgct cgtcatcaaa atcactcgca tcaaccaaac      900

```

cgttattcat tcgtgattgc gcctgagcga gacgaaatac gcgatcgctg ttaaaggac	960
aattacaac aggaatcgaa tgcaaccggc gcaggaacac tgccagcgca tcaacaatat	1020
ttcacctga atcaggatat tcttctaata cctggaatgc tgttttcccg gggatcgag	1080
tggtgagtaa ccatgcatca tcaggagtag ggataaaatg cttgatggtc ggaagggca	1140
taaatccgt cagccagttt agtctgacca tctcatctgt aacatcattg gcaacgctac	1200
ctttgccatg ttccagaac aactctggcg catcgggctt cccatacaat cgatagattg	1260
tcgcacctga ttgcccgaca ttatcgcgag cccatttata cccatataaa tcagcatcca	1320
tgttggaatt taatcgcgcg ctagagcaag acgtttcccg ttgaatatgg ctcataacac	1380
cccttgattt actgtttatg taagcagaca gttttattgt tcatgaccaa aatcccttaa	1440
cgtagctttt cgttccactg agcgtcagac cccgtagaaa agatcaagg atcttcttga	1500
gacccctttt ttctgcgctg aatctgctgc ttgcaacaa aaaaccacc gctaccagcg	1560
gtggtttgtt tgccggatca agagctacca actcttttcc cgaaggtaac tggcttcagc	1620
agagcgcaga taccaaatac tgtcttctta gtgtagccgt agtttagcca ccacttcaag	1680
aactctgtag caccgctac atacctgct ctgctaattc tgttaccagt ggctgtgtcc	1740
agtggcgata agtctgtctt taccgggttg gactcaagac gatagttacc ggataaggcg	1800
cagcggtcgg gctgaacggg ggggtcgtgc acacagccca gcttggagcg aacgacctac	1860
accgaactga gatactaca gcgtgagcta tgagaaagcg ccacgcttcc cgaagggaga	1920
aaggcggaca ggtatccggt aagcggcagg gtcggaacag gagagcgac gagggagctt	1980
ccagggggaa acgcttggtt tctttatagt cctgtcgggt ttcgccact ctgacttgag	2040
cgtagctttt tgtgatgctc gtcagggggg cggagcctat ggaaaaacgc cagcaacgcg	2100
gcctttttac ggttctctggc cttttgctgg ccttttgctc acatgttctt tctgtcgta	2160
tcccctgatt ctgtggataa ccgtattacc gcctttgagt gagctgata cgcctgccgc	2220
agccgaacga ccgagcgag cgagtcagtg agcgagggaag cggaagagcg cctgatgcgg	2280
tattttctcc ttacgcatct gtgcggtatt tcacaccgca tatatggtgc actctcagta	2340
caatctgctc tgatgccgca tagttaagcc agtatacct ccgctatgc tacgtgactg	2400
ggtcatggct gcgccccgac acccgccaac acccgctgac gcgccctgac gggcttgct	2460
gctcccgga tccgcttaca gacaagctgt gaccgtctcc gggagctgca tgtgtcagag	2520
gttttcaccg tcatcaccca aacgcgcgag gcagctgcgg taaagctcat cagcgtggtc	2580
gtgaagcgat tcacagatgt gtgcctgttc atcccgctcc agctcgttga gtttctccag	2640
aagcgtaata gtctggcttc tgataaagcg ggccatgtta agggcggttt tttctgttt	2700
ggtcactgat gcctccgtgt aaggggggatt tctgttcatt ggggtaatga taccgatgaa	2760
acgagagagg atgctcacga tacgggttac tgatgatgaa catgcccggt tactggaacg	2820



ttgtgaggg	aaacaactgg	cggatatgga	gcggcgggac	cagagaaaaa	tcactcaggg	2880
tcaatgccag	cgcttcgtta	atacagatgt	agggtgtcca	cagggtagcc	agcagcatcc	2940
tcgcatgcag	atccggaaca	taatggtgca	gggcgctgac	ttccgcgttt	ccagacttta	3000
cgaaacacgg	aaaccgaaga	ccattcatgt	tgttgctcag	gtcgcagacg	ttttgcagca	3060
gcagtcgctt	cacgttcgct	cgcgtatcgg	tgattcattc	tgctaaccag	taaggcaacc	3120
ccgccagcct	agccgggtcc	tcaacgacag	gagcacgac	atgcgcaccc	gtggggccgc	3180
catgccggcg	ataatggcct	gcttctcgcc	gaaacgtttg	gtggcgggac	cagtgcagaa	3240
ggcttgagcg	agggcgtgca	agattccgaa	taccgcaagc	gacaggccga	tcatcgtcgc	3300
gctccagcga	aagcggctct	cgccgaaaat	gaccagagc	gctgccggca	cctgtcctac	3360
gagttgcatg	ataaagaaga	cagtcataag	tcggcgacg	atagtcatgc	cccgcgcca	3420
ccggaaggag	ctgactgggt	tgaaggctct	caaggcctc	ggtcgagatc	ccggtgccta	3480
atgagtgagc	taacttacat	taattgcgtt	gcgctcactg	cccgtttcc	agtcgggaaa	3540
cctgtcgtgc	cagctgcatt	aatgaatcgg	ccaacgcgcg	gggagaggcg	gtttgcgtat	3600
tgggcgccag	ggtggttttt	cttttcacca	gtgagacggg	caacagctga	ttgcccttca	3660
ccgcctggcc	ctgagagagt	tgacgaagc	ggtccacgct	ggtttgcccc	agcagcgcaa	3720
aatcctgttt	gatggtggtt	aacggcggga	tataacatga	gctgtcttcg	gtatcgtcgt	3780
atcccaactc	cgagatatcc	gcaccaacgc	gcagcccggg	ctcggtaatg	gcgcgcattg	3840
cgccagcgcc	catctgatcg	ttggcaacca	gcacgcagct	gggaacgatg	ccctcattca	3900
gcatttgcat	ggtttgttga	aaaccggaca	tgccactcca	gtcgccctcc	cgttccgcta	3960
tcggctgaat	ttgattgcga	gtgagatatt	tatgccagcc	agccagagcg	agacgcgccg	4020
agacagaact	taatggggcc	gctaacagcg	cgatttgctg	gtgacccaat	gcgaccagat	4080
gtccacagcc	cagtcgcgta	ccgtcttcct	gggagaaaat	aatactgttg	atgggtgtct	4140
ggtcagagac	atcaagaaat	aacgccggaa	cattagtgcg	ggcagcttcc	acagcaatgg	4200
catcctggtc	atccagcgga	tagttaatga	tcagcccaat	gacgcgttgc	gcgagaagat	4260
tgtgcaccgc	cgctttacag	gcttcgacgc	cgcttcgttc	taccatcgac	accaccacgc	4320
tggcaccag	ttgatcgggc	cgagatttaa	tcgcgcgcac	aatttgcgac	ggcgcgtgca	4380
gggcccagact	ggagggtggc	acgccaatca	gcaacgactg	tttgcccgcc	agttgttgtg	4440
ccacgcgggt	gggaatgtaa	ttcagctccg	ccatcgccgc	ttccactttt	tcccgcgttt	4500
tcgcagaaac	gtggctggcc	tggttcacca	cgcgggaaac	ggtctgataa	gagacaccgg	4560
catactctgc	gacatcgat	aacgttactg	gtttcacatt	caccaccttg	aattgactct	4620
cttcggggcg	ctatcatgcc	atacccgcaa	aggtttttgcg	ccattcagatg	gtgtccggga	4680
tctcgacgct	ctcccttatg	cgactcctgc	attaggaagc	agcccagtag	taggttgagg	4740

ccgttgagca ccgccgccgc aaggaatggt gcatgcaagg agatggcgcc caacagtccc 4800  
 ccggccacgg ggccctgccac catacccacg ccgaacaacg cgctcatgag cccgaagtgg 4860  
 cgagcccgat cttcccccac ggtgatgtcg gcgatatagg cgccagcaac cgcacctgtg 4920  
 gcgccggtga tgcgggccac gatgctgccc gcgtagagga tcgagatctc gatcccgcca 4980  
 aattaatacg actcactata ggggaattgt gagcggataa caattcccct ctagaataaa 5040  
 ttttgtttaa ctttaagaag gagatataca tatggctagc atgactgggtg gacagcaaat 5100  
 gggtcgggat ccg atg agt aac ttc tct gga ttt acg aaa gga acc gat 5149  
                   Met Ser Asn Phe Ser Gly Phe Thr Lys Gly Thr Asp  
                   1                  5                  10  
 atc gca gac tta gat gcg gtg gct caa acg ctc aag aag cca gca gac 5197  
 Ile Ala Asp Leu Asp Ala Val Ala Gln Thr Leu Lys Lys Pro Ala Asp  
                   15                  20                  25  
 gat gca aac aaa gcg gtt aat gac tcg ata gca gca ttg aaa gat aag 5245  
 Asp Ala Asn Lys Ala Val Asn Asp Ser Ile Ala Ala Leu Lys Asp Lys  
                   30                  35                  40  
 cct gac aac ccg gcg cta ctt gct gac tta caa cat tca att aat aaa 5293  
 Pro Asp Asn Pro Ala Leu Leu Ala Asp Leu Gln His Ser Ile Asn Lys  
                   45                  50                  55                  60  
 tgg tcg gta att tac aat ata aac tca acc ata gtt cgt agc atg aaa 5341  
 Trp Ser Val Ile Tyr Asn Ile Asn Ser Thr Ile Val Arg Ser Met Lys  
                   65                  70                  75  
 gac tta atg caa ggc atc cta cag aag ttc cca ctc gag cac cac cac 5389  
 Asp Leu Met Gln Gly Ile Leu Gln Lys Phe Pro Leu Glu His His His  
                   80                  85                  90  
 cac cac cac tga gatccggctg ctaacaaagc ccgaaaggaa gctgagttgg 5441  
 His His His  
                   95  
 ctgctgccac cgctgagcaa taactagcat aaccccttgg ggcctctaaa cggtcttga 5501  
 ggggtttttt gctgaaagga ggaactatat ccggat 5537

<210> 14  
 <211> 5309  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> pET24b

<220>  
 <221> terminator  
 <222> (26)..(72)  
 <223> T7 terminator

<220>  
 <221> misc\_feature  
 <222> (140)..(157)  
 <223> His-tag sequence

<220>

```

<221> misc_feature
<222> (206)..(238)
<223> T7 Tag

<220>
<221> misc_feature
<222> (309)..(309)

<220>
<221> misc_feature
<222> (309)..(309)
<223> T7 transcription start

<220>
<221> promoter
<222> (310)..(327)
<223> T7 promoter

<220>
<221> misc_feature
<222> (713)..(716)
<223> Alternate start codon "gtg"

<220>
<221> CDS
<222> (713)..(1792)
<223> lacI

<220>
<221> CDS
<222> (3935)..(4747)
<223> kan

```

```

<400> 14
atccggatat agttctcct ttcagcaaaa aaccctcaa gaccggtta gaggcccaa      60
ggggttatgc tagttattgc tcagcggtagg cagcagccaa ctacgcttcc ttccgggctt    120
tggttagcagc cggatctcag tggtagtggt ggtggtgctc gagtgcgcc gcaagcttgt    180
cgacggagctc cgaattcggg tcccgaacca ttgctgtcc accagtcag ctagccatat    240
gtatatctcc ttctaaagt taaacaaat tatttctaga ggggaattgt tatccgctca    300
caattccctc atagtgcgc gtattaattt cgcgggagtc agatctcgat cctctacgcc    360
ggacgcagtc tgccggcat caccggcgcc acaggtgcgg ttgctggcgc ctatatcgcc    420
gacatcacgc atggggaaga tcgggctcgc cacttcgggc tcatgagcgc ttgtttcggc    480
gtgggtatgg tggcaggccc cgtggccggg ggactgttgg gcgccatctc cttagcatga    540
ccattccttg cggcggcggg gctcaacggc ctcaacctac tactgggctg ctctctaagt    600
caggagtcgc ataaggaga gcgtcgagat cccggacacc atcgaatggc gcaaaacctt    660
tcgcggtatg gcgatagac gcccggaaga gaggcaattc aggggtgtga at gtg aaa      718
                                     Met Lys
                                     1

cca gta acg tta tac gat gtc gca gag tat gcc ggt gtc tct tat cag      766
Pro Val Thr Leu Tyr Asp Val Ala Glu Tyr Ala Gly Val Ser Tyr Gln
      5              10              15

```

acc gtt tcc cgc gtg gtg aac cag gcc agc cac gtt tct gcg aaa acg Thr Val Ser Arg Val Val Asn Gln Ala Ser His Val Ser Ala Lys Thr	814
cgg gaa aaa gtg gaa gcg gcg atg gcg gag ctg aat tac att ccc aac Arg Glu Lys Val Glu Ala Met Ala Glu Leu Asn Tyr Ile Pro Asn	862
cgc gtg gca caa caa ctg gcg ggc aaa cag tcg ttg ctg att ggc gtt Arg Val Ala Gln Gln Leu Ala Gly Lys Gln Ser Leu Leu Ile Gly Val	910
gcc acc tcc agt ctg gcc ctg cac gcg ccg tcg caa att gtc gcg gcg Ala Thr Ser Ser Leu Ala Leu His Ala Pro Ser Gln Ile Val Ala Ala	958
att aaa tct cgc gcc gat caa ctg ggt gcc agc gtg gtg gtg tcg atg Ile Lys Ser Arg Ala Asp Gln Leu Gly Ala Ser Val Val Val Ser Met	1006
gta gaa cga agc ggc gtc gaa gcc tgt aaa gcg gcg gtg cac aat ctt Val Glu Arg Ser Gly Val Glu Ala Cys Lys Ala Val His Asn Leu	1054
ctc gcg caa cgc gtc agt ggg ctg atc att aac tat ccg ctg gat gac Leu Ala Gln Arg Val Ser Gly Leu Ile Ile Asn Tyr Pro Leu Asp Asp	1102
cag gat gcc att gct gtg gaa gct gcc tgc act aat gtt ccg gcg tta Gln Asp Ala Ile Val Glu Ala Ala Cys Thr Asn Val Pro Ala Leu	1150
ttt ctt gat gtc tct gac cag aca ccc atc aac agt att att ttc tcc Phe Leu Asp Val Ser Asp Gln Thr Pro Ile Asn Ser Ile Ile Phe Ser	1198
cat gaa gac ggt acg cga ctg gcg gtg gag cat ctg gtc gca ttg ggt His Glu Asp Gly Thr Arg Leu Gly Val Glu His Leu Val Ala Leu Gly	1246
cac cag caa atc gcg ctg tta gcg ggc cca tta agt tct gtc tcg gcg His Gln Gln Ile Ala Leu Leu Ala Gly Pro Leu Ser Ser Val Ser Ala	1294
cgt ctg cgt ctg gct ggc tgg cat aaa tat ctc act gcg aat caa att Arg Leu Arg Leu Ala Gly Trp His Lys Tyr Leu Thr Arg Asn Gln Ile	1342
cag ccg ata gcg gaa ccg gaa ggc gac tgg agt gcc atg tcc ggt ttt Gln Pro Ile Ala Arg Glu Arg Gly Asp Trp Ser Ala Met Ser Thr Phe	1390
caa caa acc atg caa atg ctg aat gag gcc atc gtt ccc act gcg atg Gln Gln Thr Met Gln Met Leu Asn Glu Gly Ile Val Pro Thr Ala Met	1438
ctg gtt gcc aac gat cag atg gcg ctg ggc gca atg gcg gcc att acc Leu Val Ala Asn Asp Gln Met Ala Leu Gly Ala Met Arg Ala Ile Thr	1486
gag tcc ggg ctg cgc gtt ggt gcg gat atc tcg gta gtg gga tac gac Glu Ser Gly Leu Arg Val Gly Ala Asp Ile Ser Val Val Gly Tyr Asp	1534

gat acc gaa gac agc tca tgt tat atc ccg ccg tta acc acc atc aaa Asp Thr Glu Asp Ser Ser Cys Tyr Ile Pro Pro Leu Thr Thr Ile Lys 275 280 285 290	1582
cag gat ttt cgc ctg ctg ggg caa acc agc gtg gac cgc ttg ctg caa Gln Asp Phe Arg Leu Leu Gly Gln Thr Ser Val Asp Arg Leu Leu Gln 295 300 305	1630
ctc tct cag ggc cag gcg gtg aag ggc aat cag ctg ttg ccc gtc tca Leu Ser Gln Gly Gln Ala Val Lys Gly Asn Gln Leu Leu Pro Val Ser 310 315 320	1678
ctg gtg aaa aga aaa acc acc ctg gcg ccc aat acg caa acc gcc tct Leu Val Lys Arg Lys Thr Thr Leu Ala Pro Asn Thr Thr Ala Ser 325 330 335	1726
ccc cgc gcg ttg gcc gat tca tta atg cag ctg gca cga cag gtt tcc Pro Arg Ala Leu Ala Asp Ser Leu Met Gln Leu Ala Arg Gln Val Ser 340 345 350	1774
cga ctg gaa agc ggg cag tgagcgcaac gcaattaatg taagttagct Arg Leu Glu Ser Gly Gln 355 360	1822
cactcattag gcaccgggat ctgcaccgat gcccttgaga gccttcaacc cagtcagctc	1882
cttcgggtgg gcgcggggca tgactatcgt cgccgcactt atgactgtct tctttatcat	1942
gcaactcgta ggacaggtgc cggcagcgct ctgggtcatt ttcggcgagg accgctttcg	2002
ctggagcgcg acgatgatcg gctgtcgct tgcggtattc ggaatcttgc acgcccctgc	2062
tcaagccttc gtcactggtc ccgccaccaa acgtttcggc gagaagcagg ccattatcgc	2122
cggcatggcg gccccacggg tgcgcatgat cgtgctcctg tcgttgagga cccggctagg	2182
ctggcggggg tgccttactg gttagcagaa tgaatcaccg atacgcgagc gaacgtgaag	2242
cgactgctgc tgcaaaacgt ctgcgacctg agcaacaaca tgaatggtct tcggtttccg	2302
tgtttcgtaa agtctggaaa cgcggaagtc agcgccctgc accattatgt tccggatctg	2362
catcgcagga tgctgtggc taccctgtgg aacacctaca tctgtattaa cgaagcgctg	2422
gcattgacct tgagtgtatt ttctctggtc ccgccgcac cataccgcca gttgtttacc	2482
ctcaacaagt tccagtaacc gggcatgttc atcatcagta acccgatcgt tgagcatcct	2542
ctctcgtttc atcggtatca ttaccccat gaacagaaat ccccttaca cggaggcatc	2602
agtgaccaa caggaaaaaa ccgcccttaa catggcccg tttatcagaa gccagacatt	2662
aacgcttctg gagaaactca acgagctgga cgcggatgaa caggcagaca tctgtgaatc	2722
gcttcacgac cacgctgatg agctttaccg cagctgcctc gcgcgtttcg gtgatgacgg	2782
tgaaaacctc tgacacatgc agctcccgga gacgggtcaca gcttgtctgt aagcggatgc	2842
cgggagcaga caagcccgtc agggcgcgtc agcgggtgtt ggcgggtgtc ggggcgcagc	2902
catgaccag tcacgtagcg atagcggagt gtatactggc ttaactatgc ggcacatgag	2962
cagattgtac tgagagtgc ccatatatgc ggtgtgaaat accgcacaga tgcgtaagga	3022

gaaaataccg catcaggcgc ttttccgctt cctcgcctac tgactcgcctg cgctcggtcg 3082  
 ttcggctgcg gcgagcgcta tcagctcact caaaggcggt aatacggtta tccacagaat 3142  
 caggggataa cgcaggaagc aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta 3202  
 aaaaaggccg gttgctggcg tttttccata ggctccgccc ccctgacgag catcacaaaa 3262  
 atcgacgctc aagtcagagg tggcgaaacc cgacaggact ataagatac caggcgtttc 3322  
 cccctggaag ctccctcgtg cgctctcctg ttccgacctt gccgcttacc gatacctgt 3382  
 ccgctttctt cccttcggga agcgtggcgc tttctcatag ctacgctgt aggtatctca 3442  
 gttcggtgta ggtcgttcgc tccaagctgg gctgtgtgca cgaaccccc gttcagcccc 3502  
 accgctgcgc cttatccggt aactatcgtc ttgagtccaa ccggttaaga cagcacttat 3562  
 cgccactggc agcagccact ggtaacagga ttacgagagc gaggtatgta ggcgggtgcta 3622  
 cagagtctct gaagtgggtg cctaactacg gctacactag aaggacagta tttgggtatct 3682  
 gcgctctgct gaagccagtt accttcggaa aaagagttgg tagctcttga tccggcaaac 3742  
 aaaccaccgc tgtagcggtt ggtttttttt ttgcaagca gcagattacg cgcagaaaaa 3802  
 aaggatctca agaagatcct ttgatctttt ctacggggtc tgacgctcag tggaacgaaa 3862  
 actcacgtta agggattttt gtcataaaca ataaaactgt ctgcttacat aaacagtaat 3922  
 acaaggggtg tt atg agc cat att caa cgg gaa acg tct tgc tct agg ccg 3973  
 Met Ser His Ile Gln Arg Glu Thr Ser Cys Ser Arg Pro  
 365 370  
 cga tta aat tcc aac atg gat gct gat tta tat ggg tat aaa tgg gct 4021  
 Arg Leu Asn Ser Asn Met Asp Ala Asp Leu Tyr Gly Tyr Lys Trp Ala  
 375 380 385  
 cgc gat aat gtc ggg caa tca ggt gcg aca atc tat cga ttg tat ggg 4069  
 Arg Asp Asn Val Gly Gln Ser Gly Ala Thr Ile Tyr Arg Leu Tyr Gly  
 390 395 400 405  
 aag ccc gat gcg cca gag ttg ttt ctg aaa cat ggc aaa ggt agc gtt 4117  
 Lys Pro Asp Ala Pro Glu Leu Phe Leu Lys His Gly Lys Gly Ser Val  
 410 415 420  
 gcc aat gat gtt aca gat gag atg gtc aga cta aac tgg ctg acg gaa 4165  
 Ala Asn Asp Val Thr Asp Glu Met Val Arg Leu Asn Trp Leu Thr Glu  
 425 430 435  
 ttt atg cct ctt ccg acc atc aag cat ttt atc cgt act cct gat gat 4213  
 Phe Met Pro Leu Pro Thr Ile Lys His Phe Ile Arg Thr Pro Asp Asp  
 440 445 450  
 gca tgg tta ctc acc act gcg atc ccc ggg aaa aca gca ttc cag gta 4261  
 Ala Trp Leu Leu Thr Thr Ala Ile Pro Gly Lys Thr Ala Phe Gln Val  
 455 460 465  
 tta gaa gaa tat cct gat tca ggt gaa aat att gtt gat gcg ctg gca 4309  
 Leu Glu Glu Tyr Pro Asp Ser Gly Glu Asn Ile Val Asp Ala Leu Ala  
 470 475 480 485  
 gtg ttc ctg gcg cgg ttg cat tcg att cct gtt tgt aat tgt cct ttt 4357  
 Val Phe Leu Arg Arg Leu His Ser Ile Pro Val Cys Asn Cys Pro Phe  
 14

490										495										500										
aac	agc	gat	cgc	gta	ttt	cgt	ctc	gct	cag	gcg	caa	tca	cga	atg	aat	4405														
Asn	Ser	Asp	Arg	Val	Phe	Arg	Leu	Ala	Gln	Ala	Gln	Ser	Arg	Met	Asn															
															505															
aac	ggg	ttg	ggt	gat	gcg	agt	gat	ttt	gat	gac	gag	cgt	aat	ggc	tgg	4453														
Asn	Gly	Leu	Val	Asp	Ala	Ser	Asp	Phe	Asp	Asp	Glu	Arg	Asn	Gly	Trp															
															520															
cct	ggt	gaa	caa	gtc	tgg	aaa	gaa	atg	cat	aaa	ctt	ttg	cca	ttc	tca	4501														
Pro	Val	Glu	Gln	Val	Trp	Lys	Glu	Met	His	Lys	Leu	Leu	Pro	Phe	Ser															
															535															
ccg	gat	tca	gtc	gtc	act	cat	ggg	gat	ttc	tca	ctt	gat	aac	ctt	att	4549														
Pro	Asp	Ser	Val	Val	Thr	His	Gly	Asp	Phe	Ser	Leu	Asp	Asn	Leu	Ile															
															550															
ttt	gac	gag	ggg	aaa	tta	ata	ggg	tgt	att	gat	ggt	gga	cga	gtc	gga	4597														
Phe	Asp	Glu	Gly	Lys	Leu	Ile	Gly	Cys	Ile	Asp	Val	Gly	Arg	Val	Gly															
															570															
atc	gca	gac	cga	tac	cag	gat	ctt	gcc	atc	cta	tgg	aac	tgc	ctc	ggg	4645														
Ile	Ala	Asp	Arg	Tyr	Gln	Asp	Leu	Ala	Ile	Leu	Trp	Asn	Cys	Leu	Gly															
															585															
gag	ttt	tct	cct	tca	tta	cag	aaa	cgg	ctt	ttt	caa	aaa	tat	ggg	att	4693														
Glu	Phe	Ser	Pro	Ser	Leu	Gln	Lys	Arg	Leu	Phe	Gln	Lys	Tyr	Gly	Ile															
															600															
gat	aat	cct	gat	atg	aat	aaa	ttg	cag	ttt	cat	ttg	atg	ctc	gat	gag	4741														
Asp	Asn	Pro	Asp	Met	Asn	Lys	Leu	Gln	Phe	His	Leu	Met	Leu	Asp	Glu															
															615															
ttt	ttc	taagaattaa	ttcatgagcg	gatacatatt	tgaatgtatt	tagaaaaata	4797																							
															630															
aacaaatagg	gggtccgcgc	acatttcccc	gaaaagtgcc	acctgaaatt	gtaaactgta	4857																								
atattttggt	aaaattcgcg	ttaaattttt	gttaaatcag	ctcatttttt	aaccaatagg	4917																								
ccgaaatcgg	caaaatccct	tataaatcaa	aagaatagac	cgagatagg	ttgagtgtg	4977																								
ttccagtttg	gaacaagagt	ccactattaa	agaacgtgga	ctccaacgtc	aaaggcgcaa	5037																								
aaaccgtcta	tcagggcgat	ggcccaactac	gtgaaccatc	accctaatac	agtttttttg	5097																								
ggtcgaggtg	ccgtaaagca	ctaaatcggg	accctaaagg	gagcccccca	tttagagctt	5157																								
gacggggaaa	gccggcgcaac	gtggcgagaa	aggaagggaa	gaaagcgaaa	ggagcgggcg	5217																								
ctagggcgct	ggcaagtgtg	gcggtcacgc	tgccgcgtaac	caccacaccc	gccgcgctta	5277																								
atgcgcgcgt	acagggcgcg	tccatttcgc	ca			5309																								

<210> 15  
 <211> 30  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> primer, HT-YscF Start

<400> 15  
cgggatccga tgagtaactt ctctggattt

30

<210> 16  
<211> 30  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> primer, HT-YscF Stop

<400> 16  
ccgctcgagt gggaacttct gtaggatgcc

30

<210> 17  
<211> 87  
<212> PRT  
<213> Yersinia pestis KIM5

<220>  
<223> YscF amino acid seuquence from Y. pestis KIM5

<220>  
<223> sequence can be found at MedLine accession number NC\_004839

<400> 17

Met Ser Asn Phe Ser Gly Phe Thr Lys Gly Thr Asp Ile Ala Asp Leu  
1 5 10 15

Asp Ala Val Ala Gln Thr Leu Lys Lys Pro Ala Asp Asp Ala Asn Lys  
20 25 30

Ala Val Asn Asp Ser Ile Ala Ala Leu Lys Asp Lys Pro Asp Asn Pro  
35 40 45

Ala Leu Leu Ala Asp Leu Gln His Ser Ile Asn Lys Trp Ser Val Ile  
50 55 60

Tyr Asn Ile Asn Ser Thr Ile Val Arg Ser Met Lys Asp Leu Met Gln  
65 70 75 80

Gly Ile Leu Gln Lys Phe Pro  
85

<210> 18  
<211> 87  
<212> PRT  
<213> Yersinia pestis C092

<220>  
<223> YscF amino acid sequence from Y. pestis C092

<220>  
<223> sequence can be found at MedLine accession number NC\_003131  
16



<400> 18

Met Ser Asn Phe Ser Gly Phe Thr Lys Gly Thr Asp Ile Ala Asp Leu  
1 5 10 15

Asp Ala Val Ala Gln Thr Leu Lys Lys Pro Ala Asp Asp Ala Asn Lys  
20 25 30

Ala Val Asn Asp Ser Ile Ala Ala Leu Lys Asp Lys Pro Asp Asn Pro  
35 40 45

Ala Leu Leu Ala Asp Leu Gln His Ser Ile Asn Lys Trp Ser Val Ile  
50 55 60

Tyr Asn Ile Asn Ser Thr Ile Val Arg Ser Met Lys Asp Leu Met Gln  
65 70 75 80

Gly Ile Leu Gln Lys Phe Pro  
85

<210> 19

<211> 87

<212> PRT

<213> Artificial Sequence

<220>

<223> Consensus sequence for YSCF

<400> 19

Met Ser Asn Phe Ser Gly Phe Thr Lys Gly Thr Asp Ile Ala Asp Leu  
1 5 10 15

Asp Ala Val Ala Gln Thr Leu Lys Lys Pro Ala Asp Asp Ala Asn Lys  
20 25 30

Ala Val Asn Asp Ser Ile Ala Ala Leu Lys Asp Lys Pro Asp Asn Pro  
35 40 45

Ala Leu Leu Ala Asp Leu Gln His Ser Ile Asn Lys Trp Ser Val Ile  
50 55 60

Tyr Asn Ile Asn Ser Thr Ile Val Arg Ser Met Lys Asp Leu Met Gln  
65 70 75 80

Gly Ile Leu Gln Lys Phe Pro  
85

<210> 20

<211> 4884

<212> DNA

<213> Artificial Sequence

<220>

<223> pBAD18-YscF

<220>

<221> misc\_feature

<222> (96)..(974)

<223> araC

<220>

<221> misc\_feature

<222> (1003)..(1020)

<223> operator O2

<220>

<221> promoter

<222> (1125)..(1153)

<223> P<sub>c</sub> promoter

<220>

<221> misc\_feature

<222> (1161)..(1182)

<223> operator O1

<220>

<221> misc\_feature

<222> (1204)..(1217)

<223> CAP site

<220>

<221> misc\_feature

<222> (1213)..(1251)

<223> operator I2 + I1

<220>

<221> promoter

<222> (1250)..(1277)

<223> P<sub>BAD</sub> promoter

<220>

<221> CDS

<222> (1323)..(1586)

<223> YscF

<220>

<221> misc\_feature

<222> (1634)..(2059)

<223> rrnB

<220>

<221> terminator

<222> (1634)..(2059)

<223> rrnB

<220>

<221> promoter

<222> (2114)..(2120)

<223> bla<sub>P3</sub> promoter

<220>

<221> misc\_feature

<222> (2152)..(3015)

<223> bla

<220>

<221> rep\_origin  
(3051)..(3509)

<223> M13 origin

<220>

<221> rep\_origin  
(3515)..(4212)

<223> pBR322 origin

<400> 20

atcgatgcat aatgtgctctg tcaaatggac gaagcagggg ttctgcaaac cctatgctac 60

tccgtcaagc cgtcaattgt ctgattcgtt accaattatg acaacttgac ggctacatca 120

ttcacttttt cttcacaacc ggcacggaac tcgctcgggc tggccccggc gcatttttta 180

aatacccgcg agaaatagag ttgatcgtca aaaccaacat tgcgaccgac ggtggcgata 240

ggcatccggg tgggtgctcaa aagcagcttc gcctggctga tacgttggtc ctgcgccag 300

cttaagacgc taatccctaa ctgctggcgg aaaagatgtg acagacgcga cggcgacaag 360

caaacatgct gtgcgacgct ggcgatatca aaattgctgt ctgcgaggtg atcgtctgatg 420

tactgacaag cctcgctac ccgattatcc atcgggtgat ggagcgactc gttaatcgct 480

tccatgcgcc gcagtaacaa ttgctcaagc agatttatcg ccagcagctc cgaatagcgc 540

ccttcccttt gcccgcgctt aatgatttgc ccaaacaggt cgctgaaatg cggctggtgc 600

gcttcatccg ggcgaagaa ccccgatttg gcaaatattg acggccagtt aagccattca 660

tgccagttag gcgcgggacg aaagtaaacc cactggtgat accattcgcg agcctccgga 720

tgacgacctg agtgatgaat ctctcctggc gggaacagca aaataccacc cggctcgcaa 780

acaaattctc gtccttgatt ttccaccacc ccttgaccgc gaatggtgag attgagaata 840

taacctttca ttcccagcgg tcggctcgata aaaaaatcga gataaccggt ggctcaatc 900

ggcggttaaac ccgccaccag atgggcatta aacgagatc ccggcagcag gggatcattt 960

tgcgcttcag ccatactttt catactcccg ccattcagag aagaaccaa ttgtccatat 1020

tgcatcagac attgccgtca ctgcgtcttt tactggctct tctcgctaac caaacggta 1080

accccgctta ttaaaagcat tctgtaacaa agcgggacca aagccatgac aaaaacgcgt 1140

aacaaaagtg tctataatca cggcagaaaa gtccacattg attatttgca cggcgctaca 1200

ctttgctatg ccatagcatt tttatccata agattagcgg atcctacctg acgcttttta 1260

tcgcaactct ctactgtttc tccatacccg tttttttggg ctagcgaatt caggaggaaa 1320

cg atg agt aac ttc tct gga ttt acg aaa gga acc gat atc gca gac 1367

Met Ser Asn Phe Ser Gly Phe Thr Lys Gly Thr Asp Ile Ala Asp  
1 5 10 15

tta gat gcg gtg gct caa acg ctc aag aag cca gca gac gat gca aac 1415

Leu Asp Ala Val Ala Gln Thr Leu Lys Lys Pro Ala Asp Asp Ala Asn  
20 25 30

aaa gcg gtt aat gac tcg ata gca gca ttg aaa gat aag cct gac aac Lys Ala Val Asn Asp Ser Ile Ala Ala Leu Lys Asp Lys Pro Asp Asn 35 40 45	1463
ccg gcg cta ctt gct gac tta caa cat tca att aat aaa tgg tcg gta Pro Ala Leu Leu Ala Asp Leu Gln His Ser Ile Asn Lys Trp Ser Val 50 55 60	1511
att tac aat ata aac tca acc ata gtt cgt agc atg aaa gac tta atg Ile Tyr Asn Ile Asn Ser Thr Ile Val Arg Ser Met Lys Asp Leu Met 65 70 75	1559
caa ggc atc cta cag aag ttc cca taa ggatccccgcg gggatcctct Gln Gly Ile Leu Gln Lys Phe Pro 80 85	1606
agagtcgacc tgcaggcatg caagcttggc tgttttggcg gatgagagaa gattttcagc	1666
ctgatacaga ttaaatcaga acgcagaagc ggtctgataa aacagaattt gcctggcggc	1726
agtagcgcgg tgggtccacc tgaccccatg ccgaactcag aagtgaacg ccgtagcgcc	1786
gatggtatgt tggggtctcc catgcgaga ttagggaaact gccaggcatc aaataaaacg	1846
aaaggctcag tcgaaagact gggcctttcg ttttatctgt tgtttgcggt tgaacgctct	1906
cctgagtagg acaaatccgc cgggagcgga ttgaaacgtt gcgaagcaac ggcccggagg	1966
gtggcgggca ggacgcccgc cataaactgc caggcatcaa attaagcaga aggccatcct	2026
gacggatggc ctttttgcgt ttctacaac tcttttgttt atttttctaa atacattcaa	2086
atatgtatcc gtcctagaga caataaccct gataaatgct tcaataatat tgaaaaaggga	2146
agagtatgag tattcaacat ttccgtgtcg cccttattcc ctttttgcg gcattttgcc	2206
ttcctgtttt tgctcaccca gaaacgctgg tgaaagtaaa agatgctgaa gatcagttgg	2266
gtgcacgagt ggggtacatc gaactggatc tcaacagcgg taagatcctt gagagttttc	2326
gccccgaaga acgttttcca atgatgagca cttttaaagt tctgctatgt ggcgcggtat	2386
tatcccggtg tgacgcgggg caagagcaac tcggtcgccg catacactat ttcagaatg	2446
acttggttga gtactacca gtcacagaaa agcatcttac ggatggcatg acagtaagag	2506
aattatgcag tgctgccata accatgagtg ataactgc ggccaactta ctctgacaa	2566
cgatcgagg accgaaggag ctaaccgctt ttttgacaa catgggggat catgtaactc	2626
gccttgatcg ttgggaaccg gagctgaatg aagccatacc aaacgacgag cgtgacacca	2686
cgatgcctcg agcaatggca acaacgttgc gcaactatt aactggcgaa ctacttactc	2746
tagcttcccg gcaacaatta atagactgga tggaggcgga taaagttgca ggaccacttc	2806
tgcgctcgcc ccttccggct ggctggttta ttgctgataa atctggagcc ggtgagcgtg	2866
ggtctcgcg tatcattgca gcaactgggc cagatggtaa gccctcccgt atcgtagtta	2926
tctacacgac ggggagtcag gcaactatgg atgaacgaaa tagacagatc gctgagatag	2986
gtgcctcact gattaagcat tggttaactgt cagaccaagt ttactcatat atactttaga	3046

ttgatttacg cgcctgtag cggcgatta agcgcggcg gtgtggtgtg tacgcgcagc	3106
gtgaccgcta cacttgccag cgcctagcg cccgctcctt tcgctttctt ccttctcttt	3166
ctcgccacgt tcgcggcgtt tccccgtaa gctctaaatc gggggctccc tttagggttc	3226
cgatttagtg ctttacgca cctcgacccc aaaaaacttg atttgggtga tggttcacgt	3286
agtgggcat cgcctgata gacggtttt cgccctttga cgttggagtc cagcttcttt	3346
aatagtgga ctttgttcca aacttgaaca acactcaacc ctatctcggt ctattctttt	3406
gatttataag ggattttgcc gatctcgcc tattggttaa aaaatgagct gatttaacaa	3466
aaatttaacg cgaattttta caaaatatta acgtttacaa tttaaaggga tctaggtgaa	3526
gatccttttt gataatctca tgacaaaat cccttaacgt gagttttcgt tccactgagc	3586
gtcagacccc gtgaaaaaga tcaaggatc ttcttgagat cctttttttc tgcgcgtaat	3646
ctgctgcttg caaacaaaa aaccaccgct accagcgggt gtttgtttgc cggatcaaga	3706
gctaccaact ctttttcga aggttaactgg ctctagcaga gcgcagatac caaatactgt	3766
ccttctagtg tagccgtagt tagggcacca cttaagaac tctgtagcac gcctacata	3826
cctcgctctg ctaatctgt taccagtggc tgctgccagt ggcgataagt cgtgtcttac	3886
cgggttgga tcaagacgat agttaccgga taaggcgag cggtcgggt gaacgggggg	3946
ttcgtgcaca cagcccagct tggagcgaac gacctacac gaactgagat acctacagcg	4006
tgagctatga gaaagcgcca cgcttcccga agggagaaag gcggacaggt atccgtaag	4066
cggcagggtc ggaacaggag agcgcacgag ggagcttcca ggggaaacg cctggtatct	4126
ttatagtctt gtcgggtttc gccacctctg acttgagcgt cgatttttgt gatgtcgtc	4186
agggggggcg agcctatgga aaaacgccag caacgcggcc tttttacggt tcttgccctt	4246
ttgtggcct tttgtcaca tgttctttcc tgcgttatcc cctgattctg tggataaccg	4306
tattaccgcc tttgagttag ctgataccgc tcgccgcagc cgaacgaccg agcgacgca	4366
gtcagtgagc gaggaagcgg aagagcgct gatcggttat ttctctta cgcatttgtg	4426
cggattttca caccgcatat ggtgcaactc cagtacaatc tgctctgatg ccgcatagtt	4486
aagccagtat acactcgtc atcgctacgt gactgggtca tggctgcgcc ccgacaccg	4546
ccaacaccg ctgacgcgcc ctgacgggt tgtctgctcc cggcatccgc ttacagacaa	4606
gctgtgaccg tctccgggag ctgcatgtgt cagaggtttt caccgtcacc accgaaacgc	4666
gcgaggcagc aaggagatgg gcgccaacag tccccggcc acggggcctg ccaccatacc	4726
cacgccgaaa caagcgctca tgagcccgaa gtggcgagcc cgattctccc catcggtgat	4786
gtcggcgata taggcgccg caaccgcacc tgtggcgccg gtgatgccg ccacgatgcg	4846
tccggcgtag aggatctgct catgtttgac agcttattc	4884